DUPLICATE ALSO

Forecasting SystemsTechnical Note No.1

Verification for Civil Aviation: Review and Proposals

By

D. J. Carrington

May 1995

Forecasting systems Division
Meteorological Office
London Road
Bracknell
Berkshire
RG12 2SZ

Permission to quote from it should be obtained from
Systems.

ORGS UKMO F

National Meteorological Library
FitzRoy Road, Exeter, Devon. EX1 3PB

DISTRIBUTION OF FS. TECHNICAL NOTES

Internal Distribution

1. The Librarian, Met. Office College, Shinfield Park, Reading.

2. Mr. A Watt, Booking - in Assistant/Molars Manager, Room E07 Met. Office Library.

3. Mr. D. Shaw, BD(FS), Room R207, Met O (CF).

5. Met. O (FS) Circulation.

· Distribution of this publication is managed by Programming Support Section of Forecasting Systems Division.
(Current occupant:- John King Tel 0344 - 854889 Internet jking @Email.Meto.Govt.UK)
2 June, 1995

METEOROLOGICAL OFFICE
LIBRARY

Summary

A review is made of the verification carried out of forecast products provided to the aviation industry. A summary of the forecast services provided to the industry is given, together with an account of the verification currently performed. Reasons why the verification should be expanded are explored, and possible areas of development are discussed in detail. Finally recommendations are made for changes to the provision of verification data.

Contents

1. Introduction

2. Forecasts for aviation
   - 2.1 Role of UKMO
   - 2.2 User groups
   - 2.3 Product groups
     - 2.3.1 Flight-planning products
     - 2.3.2 Airfield products

3. Current verification
   - 3.1 Flight-planning products
     - 3.1.1 Wind and temperature charts: verification against radiosondes and analyses
     - 3.1.2 Wind and temperature charts: use of aircraft reports
     - 3.1.3 Freezing level
     - 3.1.4 CAT
     - 3.1.5 Optimum routes
     - 3.1.6 Sig. Weather charts
   - 3.2 Airfield products

4. Reasons for development of verification
   - 4.1 Improvement of forecasts
   - 4.2 Evidence of product quality
     - 4.2.1 Contract retention
     - 4.2.2 Unverified products
     - 4.3.3 Competition with other NWP centres
   - 4.3 Analysis of forecast errors
     - 4.3.1 Marginal meteorological situations
     - 4.3.2 Response to complaints
     - 4.3.3 Monitoring of performance
   - 4.4 Responsiveness to industry developments
     - 4.4.1 Meeting industry specifications
     - 4.4.2 Changes in operators practices
     - 4.4.3 Advances in aviation technology
   - 4.5 Increases in availability of data

5. Possible expansions of verification: Upper air forecasts
   - 5.1 Current techniques
     - 5.1.1 Use of aircraft-based observations
     - 5.1.2 Further areas for verification
     - 5.1.3 Variety of grids
     - 5.1.4 Number of levels
     - 5.1.5 Frequency
     - 5.1.6 Verification of maximum wind
     - 5.1.7 Temperature forecasts for Concorde

5.2 New techniques
5.2.1 Percentage scores
5.2.2 Jet cores
5.2.3 EHW and RIEHW
5.2.4 Data for ascent/descent flight phases
5.2.5 Wind scores: additional error types
5.2.6 Clear Air Turbulence (CAT)
5.2.7 Humidity and icing
5.2.8 Cloud tops and bases
5.2.9 Cumulonimbus clouds

6. Possible expansions of verification: Airfield forecasts

6.1 Current techniques
6.1.1 Case study data
6.1.2 Use of operating categories

6.2 New techniques
6.2.1 Elimination of probability element for verification
6.2.2 Dependency on other weather types

6.3 Consistency
6.3.1 Specifications of TAFs and ICAO requirements for verification
6.3.2 Forecasters and verifiers
6.3.3 Forecasters and users

7. Dissemination of verification data

7.1 Mechanisms

7.2 Changes
7.2.1 Dissemination to customers
7.2.2 Dissemination to forecasters

8. Conclusions

8.1 High priority changes
8.1.1 Upper air forecasts
8.1.2 TAFs

8.2 Medium priority changes
8.2.1 Upper air forecasts
8.2.2 TAFs

8.3 Low priority changes
8.3.1 Upper air forecasts
8.3.2 TAFs

8.4 Longer term changes

8.5 Proposals not to be implemented

8.6 Dissemination

References

1. Introduction

The civil aviation industry is a major customer for the Meteorological Office's forecasts, and via the Civil Aviation Authority it is a major source of finance for the Office. Verification is an integral part of the forecasting process. It benefits customers by providing quantified estimates of the accuracy of forecasts they receive, and it benefits modellers and forecasters by indicating where improvements to the forecasts should be made.

There is a requirement to review whether the resources applied to verification of aviation products in particular are commensurate with the major role that the industry plays and whether current verification products meet internal and external requirements. In particular the following questions need to be addressed. Should more verification of aviation products be carried out? What verification would aviation customers like to see (if any)? Are these types of verification feasible and/or cost-effective? Are there 'specialist' areas of aviation requiring highly-tailored verification? Would such verification be justifiable? Can verification of aviation products be used more effectively within the Office to improve the forecasts provided? What attention needs to be given to the timeliness and presentation of verification data provided to the aviation community?

In this report, Section 2 sets the scene: what forecasts are currently provided and to whom. Section 3 gives a brief outline of verification done at present. Section 4 addresses reasons for developing the verification. Sections 5 and 6 detail possible developments in verification, for upper-air forecasts and for airfield forecasts respectively. In Section 7 issues related to the dissemination of verification data are discussed. Conclusions are drawn in Section 8.

2. Forecasts for aviation

2.1 Role of UKMO

The UK Meteorological Office (UKMO) is designated by WMO a *World Area Forecast Centre* (WAFC) and a Regional Area Forecast Centre (RAFC). It is in these capacities that the Office provides forecast data for aviation use. RAFCs are currently being merged with WAFCs, and the UKMO is to take on more regional responsibilities. The UKMO's status as a WAFC is not guaranteed; renewal is due in 1997. Retention of the role will be dependent upon evidence of continued good and improving performance of the Global Model and upon quality of service to the aviation industry.

2.2 User groups

The primary end-users of the UKMO's services are *commercial airlines*. They use forecasts for deciding upon flight routes and fuel loads. Within the UK the Civil Aviation Authority (*CAA*) regulates the provision of all meteorological services to airlines. Logistically airlines may receive forecast data direct from the UKMO. Financially, however, the CAA pays the UKMO for forecasts and for research and development into improvement of forecasts. Therefore the CAA also has a direct interest in the quality of the products which the UKMO provides.

National Air Traffic Services (part of the CAA) also use meteorological data. The requirements of the air traffic controllers for meteorological data are likely to increase with the future development of automated systems for controlling aircraft movements over extensive areas (e.g. the WAFTAGE system - Winds Analysed and Forecast for Tactical Aircraft Guidance over Europe).

ICAO (International Civil Aviation Authority) and *IATA* (International Air Travel Association) are not direct users of meteorological data. However, as regulatory bodies for civil aviation and for the air travel industry respectively, they have an interest in the quality of the service provided. In particular they prescribe accuracy standards which forecasts should meet.

2.3 Product groups

2.3.1 Flight-planning products

Three types of products are currently provided:

- *Gridded wind and temperature data.* Model forecast fields of wind and temperature are disseminated in GRIB (GRIdded Binary) code. Until recently the grid used was a 5° x 2° longitude-latitude grid. The standard grid required by ICAO is now an irregular latitude-longitude grid of 140km x 140km, referred to as a 'thinned' grid, which approximates to 1.25° x 1.25° at the Equator. The fields are supplied for standard pressure levels.
- Forecast *charts of wind and temperature.* These are taken from model fields and interpolated onto flight levels.
- *Significant Weather Charts.* High level charts (for above 10000') contain details of jet core strength and position, turbulence, freezing level and cloud in regions of significant turbulence. They are produced for the North Atlantic region and will also be produced for the other RAFC regions for which the UKMO has responsibility. Low level charts (for below 10000') contain significant details of wind, cloud and surface weather. They are produced for the UK. In addition to being supplied directly to customers, they also serve as guidance for forecasters in the preparation of TAFs.

The primary purposes for which these products are used are those of *route-planning*: to minimise fuel usage and flying time, and to predict fuel-load requirements as a result of the chosen route.

Further products are currently under development within NWP Division. These include products derived from model fields: forecast fields of upper wind from the Global Model are used to compute *optimum routes* for aircraft crossing the N.Atlantic by integrating the differential equation which governs the minimum time-track. The prime reason for this is to have an optimum route verification capability to assess model changes and the impact of observations.

2.3.2 Airfield products

TAFs (Terminal Aerodrome Forecasts) are routinely issued for all UK airfields. They are probability-based forecasts and include forecasts of cloud amount, cloud base, visibility and moderate or severe weather events (e.g. wind gusts, thunderstorms).

The main uses to which these forecasts are put relate to *landing conditions*: determining whether landing is likely to be possible at the destination airfield, whether mid-air 'holding' prior to landing is likely because of weather conditions, identifying the best diversion airfield, and predicting fuel-load requirements as a result of these factors. In some cases of poor weather, when a change of type of aircraft to one with more sophisticated automatic landing facilities could enable a landing to be made instead of rerouting to a different airfield, the TAF may also used for the decision of aircraft type.

3. Current verification

3.1 Flight-planning products

3.1.1 Wind and temperature fields: verification against radiosondes and analyses

Verification of model forecast fields of wind and temperature is carried out as part of the *routine verification* of the Global Model. Verification of wind and temperature for Area 2 and for the Northern Hemisphere is used as the mainstay of verification of the forecast charts of wind and temperature provided to aviation. The fields are verified on standard pressure levels. They are verified both against model analyses and against radiosonde observations. The latter have uneven spatial coverage, with notably few observations over the oceans, giving a bias in the errors to the values over land.

3.1.2 Wind and temperature fields: use of aircraft reports

Some limited verification of wind and temperature forecasts is carried out within NWP Division. Automated AMDAR, ASDAR and ACARS aircraft reports are used to calculate statistics on the accuracy of forecasts of equivalent head wind (EHW) and route integrated equivalent head wind (RIEHW) (see, for example, Forrester 1987). Statistics can be stratified in a number of different ways, for example by phase of flight (ascent, cruise, descent) or by length of sector.

3.1.3 Freezing level

Freezing levels may be derived from the model and from radiosonde profiles. A basic verification scheme, for profiles over Europe but excluding those profiles with more than one freezing level is implemented by NWP Division though *not used operationally*. Model forecast profiles are verified against radiosonde profiles.

3.1.4 CAT

Forecasts of CAT (Clear Air Turbulence) are derived operationally from Global Model forecast fields. NWP Division have developed a *pilot verification scheme* which uses measurements of CAT from ASDAR and ACARS reports.

3.1.5 Optimum routes

Optimum routes computed from forecast wind fields are not compared directly with optimum routes computed from analysis wind fields. However the time taken to fly the forecast optimum route through the analysed wind field is compared with the time taken to fly the analysed optimum route through the analysed wind field. Savings in flight time are calculated.

3.1.6 Sig. Weather charts

Verification of high level Sig. Weather charts is restricted to verification of the CAT element (see section 3.1.4). There is some overlap between the weather elements forecast for low level Sig. Weather charts and the weather events forecast in TAFs. The results of the *TAF verification* (see 3.2 below) are used as a *proxy* for low level Sig. Weather chart verification.

3.2 Airfield products

TAFs are *verified routinely against METARs* from the same airfields. TAFs are verified for 9 hours and METARs are taken half-hourly. Therefore there are several verifying METARs for each TAF. In TAFs the forecast occurrence of a weather event may be expressed either in an explicitly probabilistic way or by using terms which do not include a stated level of probability (used when events are expected to occur intermittently or temporarily or to change over a period). For verification purposes, the complete TAF is converted into an explicitly probabilistic form.

Reliability tables and reliability scores for each weather type are computed. Verification results are available for each airfield separately.

4. Reasons for development of verification

Historically a considerable amount of verification data has been generated at the UKMO but only a small proportion has been seen by customers outside the Office. Section 4.1 relates to the use of verification data internally to the UKMO, with indirect (though very important) benefits to external customers. Section 4.2 describes the role that verification data may play as a 'public relations' or sales tool and the advantages of this to the UKMO. Sections 4.3, 4.4 and 4.5 relate to ways in which external customers may benefit directly from improved provision of verification data.

4.1 Improvement of forecasts

There is an ongoing internal requirement at the UKMO to improve the forecasts. A prerequisite for improvement is to identify areas of weakness. Verification can provide a valuable diagnostic function in this respect. It can be used to locate weaknesses both in the forecast model and in manually-generated forecasts (such as TAFs).

There is scope for this verification diagnostic tool to be sharpened. For example, aircraft reports are not currently used for verification. This limits the ability to analyse and diagnose any weaknesses in the forecast model's representation of upper-level jets, which are of particular importance to aviation operations. Tenenbaum (1991) used aircraft reports to show that peak jet speeds are systematically

underestimated by the models of the UKMO and other centres. Improvement of verification by the use of aircraft-based observations would help facilitate improvement of the model.

4.2 Evidence of product quality

4.2.1 Contract retention

The environment in which the UKMO operates as a supplier of meteorological services to the aviation community is becoming increasingly competitive. The UKMO has to demonstrate the accuracy of its forecast products in order to retain its role as a WAFC, to retain its contract with the CAA, and to retain and develop its services to individual airlines. Data frequently have to be provided in such a way as to permit direct comparison with verification results from other NWP centres.

The range of fields and level of detail currently available for this purpose is rather limited. Greater detail would enable a more convincing case to be made in favour of the UKMO's services. It would also help to demonstrate that the UKMO was putting effort and resources into services for the aviation industry.

4.2.2 Unverified products

Some of the forecasts which aviation customers use are not verified, in particular Significant Weather forecasts of jet position and strength. This raises the question of whether verification data should be available for all products distributed. In practice the answer to this is that there is a customer requirement for it but with some qualification: customer interest lies not in verification of particular *forecast output* (such as specific charts), but in verification of *meteorological parameters*. Hence with regard to Sig. Weather charts there is a need to verify the forecasts of jets but not necesarily of jets as represented on a Sig. Weather chart.

4.3.3 Competition with other NWP centres

Other NWP centres, with which the UKMO competes as a provider of aviation services, are developing their verification systems. If the UKMO is to compete on the basis of the products offered as well as on the basis of the accuracy of forecasts, it needs to take note of what verification products are being developed or are already on offer from other NWP centres.

NMC have developed schemes for verifying non-continuous variables: cloud amount, cloud-top temperature, etc. (Miller & Cairns, 1993). These verification schemes have been developed primarily for use over the North American continental interior where observational data coverage is high. These schemes are more sophisticated than the verification of the TAF scheme, the nearest equivalent at the UKMO. Presumably NMC could extend their verification scheme to cover other areas, for example Northern Europe, in principle at least.

4.3 Analysis of forecast errors

4.3.1 Marginal meteorological situations

There appears to be a general appreciation amongst the aviation community of the quality of UKMO products. In the large majority of situations aviators consider forecasts to be sufficiently accurate for their requirements.

However, the occasions when forecast accuracy is most critical to them are also those occasions when there may be considerable spatial or temporal variability, such as wind strength around jet cores and low visibility at airfields. Techniques used to date for verifying upper-air forecasts (calculation of RMS and mean errors) involve spatial and/or temporal averaging and do not readily permit the accuracy of forecasts in these 'marginal' situations to be evaluated. Therefore there is a *need for verification techniques* to be developed which achieve this.

4.3.2 Response to complaints

On occasions when complaints from users of aviation forecasts are received by the UKMO, they mostly come from flight crew and relate to poor forecasts of upper-air jets. Verification of such features would enable a more informed response to be given to the complainant. Furthermore complainants would be more assured that the problems with forecasts that they encountered were understood and being addressed.

4.3.3 Monitoring of performance

The spatial and temporal averaging carried out in the calculation of verification scores has a further implication with regard to monitoring of forecast performance. Improvements (or deteriorations) in the accuracy of the model's representation of features such as jets may be masked by the averaging process. Customers are likely to be aware of any such changes, however. Therefore monitoring needs to reflect more closely the *customer's perception* of the forecasts.

4.4 Responsiveness to industry developments

The requirements of the aviation industry for verification data are not static, and the UKMO needs to be responsive to changes.

4.4.1 Meeting industry specifications

It has become apparent that the verification statistics currently computed do not enable UKMO forecasts to be evaluated in the manner required by ICAO. The Convention on International Civil Aviation (ICAO, 1992) specifies operationally desirable accuracy of forecasts, with required accuracy expressed in terms of the percentage of forecasts lying within a given range of the verifying observation. This requirement reflects a general interest in the frequency of occasions when forecasts are badly in error.

This type of assessment cannot be made using the RMS and mean errors computed for upper-air fields, neither can it be made for TAFs (for which the probability nature of the forecasts makes such a percentage approach theoretically impossible).

4.4.2 Changes in operators practices

Airline operators' use of and reliance on meteorological data has enhanced in recent years. For example airlines need to operate under tighter economic constraints than perhaps they used to. This has sharpened operators' requirements to minimise the fuel load of aircraft. The reliability of TAFs made for landing airfields needs to be known in order to make estimates of fuel requirements as reliable as possible. This has implications for flight safety as well as for financial control.

4.4.3 Advances in aviation technology

Long term plans for air traffic control involve considerable advances in the technology used. The development which would have greatest impact for meteorological purposes involves use of a computer *model* to predict the *position of aircraft*. Meteorological data would be required as input to such a model. The use by air traffic control of such a model would imply an increase in their authority in the routing of aircraft which may not prove acceptable to other sectors of the aviation industry. However, if such a development did take place, the demand for forecasts of atmospheric fields would increase substantially. Quantification of the likely accuracy would also be required, particularly of wind forecasts; this would most aptly be supplied by verification figures. These changes are several (5-10) years away.

More immediate developments involve major changes to the distribution systems of forecast data: by early 1995 the *SADIS* (SAtellite DIStribution of aviation products) system will enable forecast products to be disseminated to a large proportion of the globe by satellite. However, it does not affect requirements for verification.

4.5 Increases in availability of data

The increase in availability of observational data is not a reason *per se* for developing verification; however it has a major influence on *facilitating* that *development*. Types of observational data have become available which are not currently being used in verification. The most obvious is that of aircraft-based observations, which were discussed in section 4.1. Other categories are satellite measurements of cloud, SFLOCS lightning recorders which may be used to verify forecasts of cumulonimbus, and wind profilers.

The use of different observation types can suggest new methods of verification. For example aircraft-based observations can provide sets of measurements taken along frequently-used aircraft tracks. It is therefore possible to generate statistics for particular commercial routes as well as the traditional approach of generating statistics for defined geographical regions.

There is clearly no justification for using data for its own sake. However, if the employment in verification of forms of data not previously used can assist in meeting customer requirements, then these data should be exploited.

5. Possible expansions of verification: Upper air forecasts

5.1 Current techniques

5.1.1 Use of aircraft-based observations

There are limitations in the degree of representativeness for aviation of the verification data currently produced. These may be partly addressed by using aircraft-based observations for verification in addition to radiosonde reports. Geographical coverage by the observations would, by definition, be representative of the areas used by aviation. In particular there would be many more observations over oceanic regions, although it is recognised that the greatest density of aircraft data is over the USA. The observations are readily available on the GTS.

The implementation of such a scheme presents problems which are not present when verifying against radiosondes. Reports are valid at *asynoptic hours*. There are two possible solutions to this problem:

- use the report for verifying the forecast fields valid at the nearest synoptic hour. Any discrepancy in validity time would be ignored.
- derive a forecast value valid at the observation time by interpolation in time between two forecast fields. The error would then have to be assigned to one or other of the forecasts (preferably the closer in time) for collation of error statistics.

Reports are also made at *non-standard levels*. The best way of dealing with this would be to interpolate the observations vertically onto the standard pressure levels used for verification.

5.1.2 Further areas for verification

There may be some limited benefit in computing verification statistics for a greater number of fixed areas than Areas 2 and 5 and the Northern Hemisphere, in particular *RAFC areas*.

5.1.3 Variety of grids

There have been two developments which relate to the grids used for model fields. Firstly, model gridded forecast fields are now being disseminated on the 1.25° x 1.25° 'thinned' grid. This is now the ICAO standard grid. Secondly, aviation customers are to be offered a choice of further grids on which they receive the fields: 2.5°x1.25°, 2.5°x2.5° and 5°x2.5°. This raises an implication for verification: should verification statistics relating to fields provided to external customers be computed on the model grid, a standard verification grid (currently different from the model grid, 2.5°x2.5°), or on the grid used by the customer?

- If done on the model grid or a standard verification grid, only one set of verification statistics is required and this set is applicable to a wide audience. Comparisons with verification results from other NWP centres would be best done using results from verification on the same grid used by other centres.

- If done on the customer's grid, the verification is an exact measure of the accuracy of the fields the customer uses. Verification on different grids would produce differences in the results. These differences are likely to be small, however, and may only be seen on a level of accuracy not of interest to aviation customers. Furthermore verification on a large number of different grids would lead to a proliferation of verification results.

A distinction may be drawn between the 1.25° x 1.25° 'thinned grid' and the other grids. Due to the adoption of the *'thinned' grid* as the ICAO standard, it would seem to be important that verification data are made available as routine on this grid in addition to that done on the model grid. The other grids do not merit the same priority. It should be pointed out that a selection of fields are already verified on a 2.5°x2.5° grid as part of an exchange of verification results between NWP centres set up by the WMO Commission for Basic Systems.

5.1.4 Number of levels

Airlines are interested in having forecast winds and temperatures provided on more flight levels than at present, maybe every 2000'. The vertical resolution of the Global Model is currently insufficient for this, except by using an excessive amount of vertical interpolation. There are provisional plans to increase the number of vertical levels (from 19 to 31), which may facilitate the provision of forecasts on more levels. The other WAFC, NMC, already have 31 levels in the vertical in their model and are planning to increase the number still further.

One response, in principle at least, to this increase in the vertical resolution of the forecasts provided would be to carry out verification for *flight levels* instead of for the selection of pressure levels traditionally used for verification. This would represent a major departure from current verification practice. It would also present numerous complications: a problem of converting the pressure level of reports to actual height; interpolation of model fields to flight levels; and a decrease in the number of observations per level as the number of levels used for verification increases, which would result in less reliable statistics. A more viable option would be to increase the number of *pressure levels* for which verification data are computed. The practical upper limit on the levels used is the selection of model fields stored in the operational fieldsfiles. The additional benefit of verifying levels closer together would become more marginal as the difference between the levels decreases, since the statistics on nearby levels would be similar.

5.1.5 Frequency

More clearly of benefit would be to decrease the interval between the times for which verification is carried out. The interval is currently 12 hours. Aviation products are provided at 6-hourly intervals, however, and aviation users make decisions which are affected by changes in the situation on still shorter time-scales than that, particularly regarding the strength and position of jets. Therefore there would seem to be a clear requirement for verification products valid at *6-hourly intervals*.

5.1.6 Verification of maximum wind

The methods of verification of wind discussed so far involve combining errors for a level in the atmosphere and over a specified area; there is no direct attempt to compare the *structure* of the forecast fields with that observed. One method of verification which attempts to do this is 'maximum wind' verification. For each model grid-point in the horizontal, the vertical profiles of the forecast and analysed wind may be examined and the level in each at which the maximum wind strength occurs compared. This type of verification could be introduced within the current system and would comprise an attempt to verify the structure of wind forecasts. However, it would only be examining one dimension of the structure. A more comprehensive approach to verifying the structure is discussed in Section 5.2.2.

5.1.7 Temperature forecasts for Concorde

The requirements of Concorde for forecast data are different from those of subsonic aircraft in two distinct ways: firstly, temperature has a significantly greater effect on its operations than wind; and secondly its flight path is at higher levels (typically above 100mb). Therefore the necessary verification products are different. BA have previously expressed interest in verification data, particularly data which could demonstrate geographical variability in forecast temperature errors. The almost total lack of observational

data at the level in the atmosphere at which Concorde flies made this very difficult. In consequence, BA chose to rely on their own monitoring carried out using in-flight temperature measurements.

In the light of the interest shown in the past by BA in verification data, should data be computed especially for Concorde? The best approach would probably be to *reconsider* this question *once aircraft observations are used* for verification (see 5.1.1). Concorde's temperature observations are used for data assimilation in the Global Model and would therefore be available for verification. With sufficient observational data, computing verification statistics for Concorde routinely would become viable.

5.2 New techniques

5.2.1 Percentage scores

The computation of 'percentage scores' (see 4.4.1 above) in addition to RMS and mean errors would be fairly *straightforward* for upper-air forecasts of wind and temperature. Threshold values for the difference between forecasts and observations have already been stipulated by ICAO (1992).

5.2.2 Jet cores

Verification of the parts of upper-air wind forecasts of greatest interest where there are considerable spatial or temporal changes (see 4.3.1 above) requires specific verification of jet cores.

Verification of a forecast of a jet is *problematical*. Firstly, does just the jet maximum need to be verified or a more extensive part of the jet? If just the jet maximum is of interest, then verification of the maximum speed is relatively straightforward. Verification of position would require measurement of both horizontal and vertical displacement. Accuracy in jet position transverse to the direction of the jet is more critical than accuracy longitudinal to the flow, and this would have to be reflected in the method of assessment. If a more extensive part than the core of the jet is required for verification, then the problem is more complex. Some threshold speed for the 'jet core' could be defined. The boundary to this core would be a surface. Verifying this surface would be very complex. An approximation could be made by just verifying the strongest part of the jet; this would effectively involve reverting to verifying the jet maximum as described above.

A jet core is a continuous feature, and so effective verification would require use of a continuous field for verification. Observations would not provide continuous enough coverage for this purpose, and so the verification would need to be carried out *against analyses*.

5.2.3 EHW and RIEHW

The categorisation of verification statistics by pressure level and length of forecast presents difficulties for the operational aviator. He needs to estimate the accuracy of the prediction of the time to be taken by a flight; a flight will span several flight levels (including those for the ascent and descent phases), and a long-haul flight will additionally span several verifying times. The statistics the operator wants would be a weighted mean of statistics for these levels and times. RIEHW statistics automatically incorporate these weights. Verification of RIEHW against aircraft data is an international standard agreed by WMO.

The problem can be tackled on two levels. On the more sophisticated level, statistics can be generated for principal city-to-city routes. A simpler approach is to compile statistics for frequently-used flight paths common to several routes, for example between 10°W and 60°W across the North Atlantic.

These data would provide a measure of forecast accuracy *complementary to mean and RMS scores*. Some airlines already compile data in this way, incorporating factors which affect flight time additional to meteorological ones. For these data to be available as a product from the UKMO may be welcomed by airlines, especially if the facility to use them to compare the performance of UKMO with that of other NWP centres was also available.

5.2.4 Data for ascent/descent flight phases

If aircraft-based observations taken during the ascent and descent phases of flights are used for verification, error statistics for the lower atmosphere may be categorised by ascent/descent instead of (or as well as) by pressure level. Verification is already carried out on this basis in NWP Division. There is not a requirement from users for such data at the moment. However, it will be a *longer term requirement* if aircraft position models come into use.

5.2.5 Wind scores: additional error types

RMS vector wind errors and mean speed errors constitute the only measures of upper wind errors used by aviation customers. A range of other scores are currently computed, such as RMS speed errors and mean and RMS direction errors. There are further scores which could also be developed, such as a *3-way correlation coefficient* to correlate the speed of the forecast and of the analysis and the directional error. This may prove useful to modellers. It does not appear likely that there would be a requirement from aviation customers for any further scores.

5.2.6 Clear Air Turbulence (CAT)

Verification of the CAT forecasts provided on Sig. Weather charts cannot be contemplated until there are *more observations* from aircraft fitted with turbulence instruments. A scheme would need to integrate the frequency of occurrence of turbulence in order to verify the forecasts which are probability-based.

5.2.7 Humidity and icing

If an air traffic control model to predict aircraft positions to a high level of accuracy comes into use (see 4.4.3), aircraft icing would need to be included in the model since it slows aircraft down. Verification of the humidity fields as well as of the temperature fields would be necessary. This is not a requirement for the near future, but will probably be a *longer-term requirement*, however.

5.2.8 Cloud tops and bases

Forecasts of cloud top and cloud base are included in low level Sig. Weather charts. NMC carry out some verification of these parameters (Miller & Cairns, 1993): should the UKMO do likewise? NMC carries out this verification over the USA. TAFs as well as Sig. Weather charts contain forecasts of cloud base. These are verified and these verification results are taken as providing an *adequate estimate* of the accuracy of Sig. Weather cloud base forecasts.

In principle, the methodology used for verification of the TAF forecasts of cloud base could be extended to the cloud base element in low level Sig. Weather forecasts. A contingency table based approach would be the most suitable. Verification could only be carried out for cases when both cloud is forecast and it occurs.

Verification of cloud tops would present a very difficult problem, since no direct observations are reported. The approach adopted by NMC is to take satellite estimates of cloud-top temperature and obtain the height at which this temperature occurs in radiosonde profiles. There are clearly limitations to this method; for example, it will be more effective when cloud cover is continuous than when it is broken.

5.2.9 Cumulonimbus clouds

Cumulonimbus clouds present a potential hazard to aviation, and if they are forecast they are included on Sig. Weather charts. Their dispersed nature and the difficulty in identifying them reliably makes verification problematical. An indirect approach which could be taken would be to use data from SFLOCS *lightning recorders*. The existence of lightning incidents within a given area over a time period could be taken as being indicative of cumulonimbus. However, in relation to other areas of verification this should not be a high priority.

6. Possible expansions of verification: Airfield forecasts

6.1 Current techniques

6.1.1 Case study data

Most of the elements in a TAF forecast which are of the type which either occur or don't occur (e.g. thunderstorms) are climatologically rare. The users of the forecasts are primarily interested in those occasions when these events are forecast and/or occur (as discussed in section 4.3.1). Therefore, in addition to the calculation of verification scores which summarise all forecasts, there is a need for further verification which *concentrates on these occasions*. Because TAF forecasts are probability based, an individual forecast cannot be described as correct or incorrect; a number of events must be combined statistically in order to draw reliable conclusions. Using the current scheme the scores calculated go as far as they can in isolating these 'marginal' occasions.

A similar situation arises for weather types which can take a continuous range of values, notably cloud base and visibility. Errors are *more critical when weather conditions are bad*. For example, a forecast visibility error of 100m is more significant when the visibility is 200m than when it is 2000m. The current scheme partly makes allowance for this by using a low value threshold, for example 350m for visibility. However, the scheme has its failings: if the forecast is for 400m and the actual is 300m, then the event is recorded as missed by the forecast; if however the forecast is for 300m and the actual is 100m, the verification scheme scores this as a successful forecast even though the error is both larger in magnitude and potentially more critical in its effect.

If TAFs were not probability based, then one of several approaches could be taken to solve this problem for continuous weather types; for example an RMS error could be computed for those occasions when either the forecast or observed value is below a certain threshold, or contingency tables could be used. However, there is no proven verification method which takes into account both a variable value of the forecast/observed event and probability in the forecast: probability-based methods assess the forecast of events which may be defined as having either occurred or not occurred.

For both occurence/non-occurence and continuous weather types, there is a straightforward solution: to make the raw forecast and observational data available to customers, effectively on a *case-study* basis. This does not provide a very satisfactory answer, because summary statistics are more helpful to a customer than large volumes of detailed data. The problem will be taken up again in section 6.2.1.

6.1.2 Use of operating categories

TAFs are used to allocate an operating category to an airfield (category I, II, IIIa or IIIb). These categories are defined in terms of minimum visibility and/or cloud base. Particular aircraft are able to operate within certain categories, depending primarily on the level of autopilot facilities with which the aircraft is equipped. Therefore the decision of an airline as to which aircraft to use may be affected by the operating category selected.

In principle, these operating category *thresholds* provide the most natural values to use in verification. However, these thresholds differ between aircraft types. So whilst the thresholds used in verification should be chosen so as reflect the operating methods of the aviation industry in general, it does not seem viable to tie them tightly to particular aircraft categories.

6.2 New techniques

6.2.1 Elimination of probability element for verification

One approach to dealing with the difficulty in verifying probability forecasts has been devised by Gordon (1993). Forecasts and observations (SPECIs as well as METARs) are grouped into a time-window of three hours; it is assumed that the worst conditions forecast with non-zero probability in that period (i.e. the lowest cloud base and visibility and the highest wind speed) should be observed at least once in that period

for it to have been included in the forecast. Forecasts are then verified on a correct or incorrect basis and contingency tables can be constructed.

Because of the prepocessing done, the verification is effectively of a modified form of the TAF. For the forecaster this verification does not provide a clear measure of the accuracy of the forecasts as he gives them. However, the aviation customer can be given a clear answer to the type of question in which he is most interested: what proportion of forecasts of visibility (for example) below a certain threshold are correct? This approach seems to offer a *viable method* of both negotiating the probability problem and meeting customer requirements, though the details of such a scheme would have to be discussed with customers. The only alternative would seem to be the lengthy process of examining individual forecasts on a case-study basis (as in section 6.1.1).

6.2.2 Dependency on other weather types

For forecasts of wind, a further issue may be addressed. The sensitivity of aircraft operations to forecast errors depends on the *wind direction*: crosswind strength is more important than head/tailwind strength. Verification data which made allowance for this would be of value to the airlines. Knowledge of runway direction at each airport would be required.

6.3 Consistency

There are further issues relating to the way in which TAFs are handled which do not involve expansion of verification so much as changing and tightening current practice. There is at present a lack of consistency in three areas.

6.3.1 Specifications of TAFs and ICAO requirements for verification

ICAO's requirements for verification of airfield forecasts (ICAO, 1992) stipulate that a given percentage of forecasts should lie within a stated accuracy range. So long as TAFs are probability-based, this requirement *cannot be met*. For consistency to be reached between ICAO's requirements for verification and the provision of verification by the UKMO (and other suppliers of TAFs), either the ICAO requirements must be amended or TAFs must be issued in a non-probability-based form.

6.3.2 Forecasters and verifiers

The precise meaning of the terms TEMPO, PROB and GRADU in TAFs may be different between that used by *forecasters* and that used in *verification*, particularly when the terms are used in combination. For verification all forecasts are converted into probabilities. When the verification scheme was written, these probabilities were designed to reflect the meaning of the forecasts. However, changes in TAF specifications and in forecaster practices have probably led to some divergence between the forecasters' intended meaning and the meaning interpreted by the verification.

6.3.3 Forecasters and users

TAFs are used for decisions concerning fuel loading and choice of landing aerodrome. The decision-maker must commit himself to one of the options available. A *probability-based forecast* is therefore used to make a *non-probability-based decision*. The aviation industry has operational guidelines as to what forecast probability thresholds should be taken to imply an event is definitely (or definitely not) going to occur. This may not reflect correctly the forecaster's understanding of the likelihood of an occurrence. Whilst this discrepancy is an issue to be resolved between forecasters and users, it is worthy of note here since it relates to how users will interpret verification figures of the forecasts.

7. Dissemination of verification data

7.1 Mechanisms

The International Services (IS) group in FS Division is in direct contact with the CAA and with airlines as a natural part of its operations. Liaison regarding the provision and improvement of services is carried out by IS. It is therefore appropriate that, for the most part, verification data for the aviation industry continue to be *disseminated through IS.*

For IS to be the main mediary for verification data has the following implications: firstly that IS must keep abreast of the types of verification data generate by FS4 and appropriate to the aviation industry; and secondly that IS either obtain recent verification results on a regular basis or have ready access to them.

Negotiation and arrangement of contracts with the aviation industry is carried out by Head of Aviation Services in the Sales and Business Directorate. When changes to the verification data provided are to be made, direct contact between FS4 and aviation customers will be necessary in order that user requirements may be mutually agreed. This contact should be made in liaison with Head of Aviation Services.

7.2 Changes

7.2.1 Dissemination to customers

Changes in the need for dissemination of verification data will *arise naturally out of implementations* of any the proposals outlined in sections 5 & 6. The driving force behind most of these proposals is the ambition of providing a better service to external customers. It is therefore vital that extra verification data generated at the UKMO reaches those customers. In particular greater emphasis needs to be given to the dissemination of verification data to end-users of forecasts, namely route-planners, flight crew and air traffic controllers.

7.2.2 Dissemination to forecasters

Forecasters receive verification statistics for TAF forecasts on a monthly basis. There is little provision of quantitative estimates of forecast accuracy in 'real time'. For this to be done a locally based system would be required, probably using a PC. Systems such as this have been developed elsewhere, for example in the USA at Dallas Fort Worth Air Traffic Control Center.

For such a system to be efficient it would need to span several (if not all) types of forecast produced, requiring a unified system of software and of data input and processing. In particular public forecasts as well as forecasts specifically for aviation should be incorporated. If the system were locally based it could be extended to enable comparison of the performance of different forecasters as well as to measure the improvements made by manual forecaster input. For these reasons this issue needs to be considered in a broader context than that of aviation.

8. Conclusions

Significant changes have taken place in recent years which affect the provision of forecasts to the aviation industry. Competition between National Meteorological Services for the supply of meteorological data to the aviation industry has increased in recent years. Technology used within the industry has advanced, and the need for accurate forecasts has grown. Furthermore an increasingly 'customer-oriented' culture has developed. These trends are likely to continue.

The verification service has to be responsive to these trends. A number of possible improvements have been outlined in this paper. More ideas for improvements have been proposed than are likely to be implemented in the light of the resources probably available. Some are more feasible than others to implement, some require more urgent action than others. Therefore recommendations are made below of the priority to be given to the proposals. Recognition is given to the fact that some proposals are not appropriate at present but may be important in the longer term. There are some proposals which are recommended not to be quantitative implemented.

8.1 High priority changes

None of these changes require the development of substantially new verification methods.

8.1.1 Upper air forecasts

1. There is a clear requirement for upper air forecasts of wind and temperature to be verified in terms of the *percentage frequency* of occurrences when errors are within specified bounds (ICAO, 1992).
2. Use should be made of *aircraft observations*. The advantage of areal coverage that they would bring would enable, for the first time, verification statistics to be representative of the regions (including the oceans) for which forecasts are used in aviation.

8.1.2 TAFs

1. In addition to the summary statistics currently produced, a system should be set up so that forecast and observation data can be made available for *significant events* of interest to aviation users.
2. The *inconsistencies* in the way that forecasters, verifiers and users interpret TAFs must be addressed and eliminated.

8.2 Medium priority changes

8.2.1 Upper air forecasts

1. To reflect the way forecasts are used operationally by the aviation industry, the frequency of validity of verification data should be increased from 12-hourly to *6-hourly*. This would require relatively little change to the verification system.
2. Statistics over *RAFC areas* should be computed so that verification data can be 'packaged' according to area using the same method of geographical partitioning as that used for the forecasts.
3. A version of the scheme devised in NWP Division to compute and verify *equivalent headwinds* and route integrated equivalent headwinds should be implemented operationally.

8.2.2 TAFs

1. A scheme should be set up in which the probability element of TAFs is converted into a *non-probabilistic* form. This will enable contingency tables of forecast and observed values to be constructed for all weather variables.

8.3 Low priority changes

8.3.1 Upper air forecasts

1. An increase in the *number of levels* used for verification may eventually be worthwhile. A small number of additional levels should be added at first. Only if errors are significantly different on these levels should the number of levels be increased further. For diagnostic purposes, further benefit could be derived from data for the existing levels before study of more detailed variation with height becomes appropriate.

2. Verification of *jet cores*, including both strength and position of the jets, would require substantial development. The main interested parties external to the Office would be flight crew and not CAA or airlines' management. This alone would probably not justify such an investment. However, it would also provide a valuable tool for model diagnosis and development.
3. There are insufficient observations of cloud base over the oceans for a successful verification scheme to be implemented. However, for the RAFC areas primarily over land, statistics of *cloud base* would provide a valuable addition to the verification data available for customers. A version of the TAF scheme could be implemented for Europe.
4. Once aircraft reports are used in the operational verification system, it should be ascertained whether British Airways would require regular provision of verification data for the temperature forecasts made for *Concorde*.
5. The feasibility of using lightning reports to verify forecasts of cumulonimbus should be investigated.

8.3.2 TAFs
1. Consideration should be given to developing new statistics for visibility and cloudbase which give explicit emphasis to the requirement for *greater accuracy* in *severe weather* conditions.
2. A verification statistic should be developed for wind which makes allowance for the dependency of aircraft operations on the *wind direction*. A poorer score would be given to forecast errors in crosswinds than to errors of the same magnitude in head/tailwinds.

8.4 Longer term changes

If the automation of air traffic control procedes apace and detailed forecasts are used to predict aircraft positions, estimates of forecast accuracy would be a necessary input to the air traffic control models. Verification data would be required on a very detailed scale. Some extensions to the verification which are not required at present, including those developed in NWP Division, would be warranted.

1. Use of verification data from flights in their *ascent and descent stages* should be used, since they provide evidence of the spatial variablity of forecast errors, particularly in the vicinity of airports.
2. *Relative humidity* fields should be verified since forecasts of relative humidity would be used to predict aircraft icing.
3. Provided that the volume of aircraft-based observational data becomes sufficient, forecasts of *Clear Air Turbulence* should be verified operationally.

8.5 Proposals not to be implemented

1. Computation of verification statistics *on all the grids* on which forecasts are disseminated should not be implemented. The benefit to aviation customers would be slight, and the proliferation of datasets (which would be nearly identical) and the computation required for their calculation would not justify such a small return.
2. There is no foreseen requirement for *error types for verifying upper-air* wind forecasts further to those currently produced and the percentage scores detailed in 8.1.1.
3. *'Maximum wind'* verification would provide an extremely limited perspective on verification of jets because it is only one-dimensional. The more realistic verification method of jet cores recommended in 8.2.1 would be strongly preferred.
4. *Cloud top verification* would require too great an investment in resources to be worthwhile.

8.6 Dissemination

The International Services Group in FS Division should be the main vehicle for dissemination of verification data.

References

Forrester D A (1987). Monitoring of equivalent headwind and temperature forecasts from the 15-level coarse mesh model. S.I. Tech. Note No. 49, UK Met. Office.

Gordon N (1993). Verification of terminal forecasts. *Proceedings Fifth International Conference on Aviation Weather Systems*, 2-6 Aug. 1993, Vienna, Virginia.

ICAO (1992). The Meteorological Service for International Air Navigation. (Annex 3 to the Convention on International Civil Aviation), Eleventh Edition, July 1992.

Miller R J & Cairns M M (1993). A preliminary evaluation of numerical analysis and prediction of aviation-impact variables. *Proceedings Fifth International Conference on Aviation Weather Systems*, 2-6 Aug. 1993, Vienna, Virginia.

Tenenbaum J (1991). Jet stream winds: comparisons of analyses with independent aircraft data over southwest Asia. *Wea and Forecasting 6*, 320-336.

FORECASTING SYSTEMS TECHNICAL NOTE

NUMBER	TITLE	AUTHOR	DATE
1	Verification for civil aviation Review & Proposal suite.	D. J. Carrington	1995